Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Functional Analysis 236 (2006) 244–264

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Gradient estimates for diffusion semigroups with singular coefficients ☆

Enrico Priola [a], Feng-Yu Wang [b,*]

[a] *Department of Mathematics, University of Torino, Torino 10138, Italy*

[b] *School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China*

Received 9 March 2005; accepted 14 December 2005

Communicated by L. Gross

Abstract

Uniform gradient estimates are derived for diffusion semigroups, possibly with potential, generated by second order elliptic operators having irregular and unbounded coefficients. We first consider the $\mathbb{R}^d$-case, by using the coupling method. Due to the singularity of the coefficients, the coupling process we construct is not strongly Markovian, so that additional difficulties arise in the study. Then, more generally, we treat the case of a possibly unbounded smooth domain of $\mathbb{R}^d$ with Dirichlet boundary conditions. We stress that the resulting estimates are new even in the $\mathbb{R}^d$-case and that the coefficients can be Hölder continuous. Our results also imply a new Liouville theorem for space–time bounded harmonic functions with respect to the underlying diffusion semigroup.

Keywords: Gradient estimates; Diffusion semigroups; Coupling

1. Introduction

In this paper we establish uniform gradient estimates for Schrödinger diffusion semigroups generated by second order elliptic differential operators L_0 having irregular and possibly unbounded coefficients. We first consider L_0 on $\mathbb{R}^d$ and then, more generally, on a possibly

☆ Supported in part by NNSFC(10121101) and RFDP, and by the Italian Murst project "Equazioni di Kolmogorov."

* Corresponding author.

E-mail addresses: priola@dm.unito.it (E. Priola), wangfy@bnu.edu.cn (F.-Y. Wang).

0022-1236/$ – see front matter

doi:10.1016/j.jfa.2005.12.010

unbounded smooth closed domain $D \subset \mathbb{R}^d$ with Dirichlet boundary conditions. The operator L_0 can be written as

$$L_0 = \frac{1}{2} \sum_{i,j=1}^{d} q_{ij} \partial_i \partial_j + \sum_{i=1}^{d} b_i \partial_i - V =: L - V.$$

The coefficients q_{ij} and b_i are assumed to be at least continuous on D; the $d \times d$ matrix $q(x) = (q_{ij}(x))$ is uniformly positive definite and the potential term V is nonnegative and Borel on D. Our precise assumptions are collected in Hypothesis 4.1 (and in Hypothesis 3.1 when $D = \mathbb{R}^d$). We stress that, in particular, we are able to treat the case of Hölder continuous coefficients q and b. This seems to be the first paper that establishes uniform gradient estimates for diffusion semigroups when the coefficients are unbounded and not assumed to be locally Lipschitz.

It is well known that, under suitable conditions on the growth of the coefficients, the L-diffusion process uniquely exists up to the boundary ∂D of D (see [12,20,22]). Define the Schrödinger diffusion semigroup

$$P_t^D f(x) := \mathbb{E}^x \big(f(x_t) \mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s} 1_{\{\tau > t\}} \big), \quad x \in D,\ f \in \mathcal{B}_b^+(D),$$

where $\mathbb{E}^x$ is the expectation taken with respect to the distribution $\mathbb{P}^x$ of the L-diffusion process starting from x, x_t denotes the canonical process on a probability space of continuous trajectories with values in D, and

$$\tau := \inf\{t \geqslant 0\colon\ x_t \in \partial D\}.$$

Moreover, $\mathcal{B}_b^+(D)$ is the set of all bounded nonnegative Borel real functions on D. If f is regular enough and q, b and V are at least locally Hölder continuous, then it is known that $P_t^D f$ gives the classical bounded solution to the parabolic Dirichlet problem involving L_0, see, for instance, [19,20]. In this paper we will establish gradient inequalities of the type

$$\big\| \nabla P_t^D f \big\|_\infty := \sup_{x \in D} \big| \nabla P_t^D f(x) \big| \leqslant \frac{c}{\sqrt{t \wedge 1}} \| f \|_\infty, \quad t > 0,\ f \in \mathcal{B}_b^+(D), \tag{1.1}$$

involving the supremum norm $\| \cdot \|_\infty$. Note that, since we do not assume that q, b and V are locally Hölder continuous, $P_t^D f$ could be nondifferentiable in x; hence we consider a natural generalization of $|\nabla P_t^D f(x)|$, see (2.3). In addition to (1.1) we prove that, even for non-convex domains D, the map $x \mapsto P_t^D f(x)$ is globally Lipschitz continuous up to the boundary of D, for any $t > 0$.

We stress that uniform gradient estimates like (1.1) also imply the equivalence between functional inequalities and corresponding isoperimetric inequalities, see, e.g., [14,25] for details.

If the coefficients q, b and V are globally bounded and uniformly continuous on D, then L_0 generates an analytic semigroup on the space $UC_b(D)$ of all uniformly continuous and bounded functions on D, endowed with the supremum norm, see [17] and references therein. As a consequence of this nontrivial generation result, one obtains the gradient estimates (1.1). On the other hand, when the coefficients are unbounded, generation of analytic semigroups is no longer true in general even if the coefficients are assumed to be very regular, see [8]. Recently, uniform gradient estimates have been intensively investigated when the coefficients q, b and V are unbounded but at least of class C^1. To this purpose, both analytic methods, see, for instance, [1–3,10,18], and probabilistic arguments, see, for instance, [4,6,21,23,25], have been used on domains of $\mathbb{R}^d$ and on manifolds. Even if the coefficients are regular, uniform gradient estimates are not always true. Counterexamples are given in [1], if $D = \mathbb{R}^d$, and in [25] for domains of $\mathbb{R}^d$. On the other hand, when $D \neq \mathbb{R}^d$, uniform gradient estimates hold under an additional assumption on the boundary ∂D, which has been introduced in [10].

To prove (1.1), the main technique we adopt here is the coupling method as in [6,23,25]. This method is in contrast with the analytic approach mainly based on a priori estimates and the maximum principle. Our coupling is constructed in the following way: by reflection for a constant part $\lambda_0 I$ and by parallel displacement (called "March Coupling" in [5]) for the remaining part. Since our coefficients are only continuous and the coupling elliptic operator is degenerate, we have additional technical difficulties in proving the existence of the coupling process such that its two marginal processes move together after the coupling time (namely, the first time they meet). This is why we regularize also the coefficients of the coupling operator, introducing an additional term, depending on $\varepsilon > 0$; in this way the coefficients of the coupling operators are continuous on the whole space, rather than merely well defined outside the diagonal as it is usual (cf. [5,15]). We only prove the existence of such a family of coupling processes, depending on ε, but not their uniqueness as it happens in other cases of coupling available in the literature (see, e.g., [15]). This lack of uniqueness implies also that we cannot use the strong Markov property for the coupling process, see Section 3.1 and the proof of Theorem 4.5. Nevertheless, under reasonable conditions, our coupling will finally lead to non-trivial gradient estimates for P_t^D when $\varepsilon \to 0^+$. We also mention that our gradient estimates lead naturally to a new Liouville theorem for space–time bounded harmonic functions (see Theorem 3.6).

After some preliminaries given in Section 2, we study the global case (i.e., $D = \mathbb{R}^d$) in Section 3. Then we extend the resulting estimates to general smooth domains in Section 4. In this section we consider an assumption on the boundary ∂D which generalizes the one introduced in [10], see also Example 4.3. Concerning the operator L_0, if there is no potential term V, we substantially extend the assumptions in [10], both with respect to the regularity of the coefficients q and b, as already mentioned, and with respect to their growth at infinity. If $V \neq 0$, then our assumption of boundedness and measurability of V is different and not comparable with the corresponding one in [10] (see also Corollary 3.5 which deals with operators L_0 having Lipschitz potentials V).

Finally, we remark that our uniform gradient estimates hold also for Neumann semigroups on smooth convex domains as soon as the coupling processes with reflecting boundary exist. Indeed, according to [24], the convexity of D would help the coupling to be successful and would allow one to get the uniform gradient estimates.

2. Preliminaries and basic notation

The inner product and the Euclidean norm in $\mathbb{R}^d$ will be denoted by $\langle\cdot,\cdot\rangle$ and $|\cdot|$. Moreover, we set

$$\rho(x, y) := |x - y|, \quad x, y \in \mathbb{R}^d.$$

The space of all $r \times d$ matrices will be denoted by $\mathcal{L}(\mathbb{R}^r, \mathbb{R}^d)$; further $\mathcal{L}(\mathbb{R}^d, \mathbb{R}^d) = \mathcal{L}(\mathbb{R}^d)$. If $A \in \mathcal{L}(\mathbb{R}^r, \mathbb{R}^d)$, its adjoint is denoted by $A^* \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^r)$. For any $h \in \mathbb{R}^d$, we consider

$hh^* \in \mathcal{L}(\mathbb{R}^d)$, $hh^*(k) := \langle h, k\rangle h$, for any $k \in \mathbb{R}^d$. Moreover, we set

$$\|A\| := \big(\mathrm{tr}(AA^*)\big)^{1/2}, \quad A \in \mathcal{L}\big(\mathbb{R}^r, \mathbb{R}^d\big).$$

By D we shall denote a possibly unbounded *closed* connected subset of $\mathbb{R}^d$ with smooth boundary ∂D, interior $\mathring{D}$ and such that the *distance function*

$$\rho_{\partial D}(x) := \inf_{y \in \partial D} |x - y|, \quad x \in \mathbb{R}^d, \tag{2.1}$$

is a C^2-function with bounded second order derivatives in D_α, for some $\alpha > 0$, where

$$D_\alpha = \big\{x \in \mathbb{R}^d\colon\ \rho_{\partial D}(x) \leqslant \alpha\big\}. \tag{2.2}$$

It will be enough that ∂D is a uniformly C^2-boundary, see, for instance, [11] or [10, Appendix] for more details. Of course D can be the whole $\mathbb{R}^d$. Moreover, $\rho_{\partial D}$ is Lipschitz continuous with $|\nabla \rho_{\partial D}(x)| = 1$, $x \in D_\alpha$.

Let $f : D \to \mathbb{R}$ be a locally Lipschitz function, we set

$$\big|\nabla f(x)\big| := \limsup_{y \to x} \frac{|f(x) - f(y)|}{|x - y|}, \tag{2.3}$$

for any $x \in D$. Note that if f is differentiable in $x \in \mathring{D}$, then $|\nabla f(x)|$ is just the norm of the gradient of f in x.

Let us fix some notation on the martingale problem and coupling for diffusions, see also [5, 12,16,20,22]. Let $\Omega^d := C([0, \infty); \mathbb{R}^d)$ be endowed with the metric of the uniform convergence on bounded intervals. Let x_t be the canonical process on Ω^d, i.e., $x_t(\omega) = \omega(t)$, $\omega \in \Omega^d$, $t \geqslant 0$. We consider on Ω^d the natural filtration $(\mathcal{F}_t)_{t \geqslant 0}$ associated to x_t. Note that $\sigma(\mathcal{F}_t,\ t \geqslant 0)$ is the Borel σ-algebra of Ω^d.

Recall that a *martingale solution for L* is a family of probability measures $\mathbb{P}^x$, $x \in \mathbb{R}^d$, on Ω^d such that $\mathbb{P}^x(\omega(0) = x) = 1$ and for each $f \in C_0^2(\mathbb{R}^d)$,

$$M_t^f := f(x_t) - \int_0^t Lf(x_s)\,\mathrm{d}s \tag{2.4}$$

is an $\mathcal{F}_t$-martingale (recall that if $S \subset \mathbb{R}^d$, $C_0^2(S)$ denotes the space of all C^2-functions $f : S \to \mathbb{R}$ with compact support). A *coupling distribution* for $\mathbb{P}^x$ and $\mathbb{P}^y$, $x, y \in \mathbb{R}^d$, is a probability measure $\mathbb{P}^{x,y}$ on $\Omega^{2d} = (\Omega^d)^2$ such that the first and the second marginal distributions of $\mathbb{P}^{x,y}$ are just $\mathbb{P}^x$ and $\mathbb{P}^y$, respectively. That is,

$$\mathbb{P}^x = \mathbb{P}^{x,y} \circ \pi_1^{-1}, \quad \mathbb{P}^y = \mathbb{P}^{x,y} \circ \pi_2^{-1}, \quad x, y \in \mathbb{R}^d, \tag{2.5}$$

where $\pi_i : (x_1, x_2) \mapsto x_i$, $i = 1, 2$. We finally define the *coupling time*:

$$T := \inf\{t \geqslant 0\colon\ x_t = y_t\}, \tag{2.6}$$

where we set $\inf \emptyset = \infty$ by convention.

In Section 4, we will also consider martingale problems on domains, see [20, Sections 1.12 and 1.13] for more details. Let $\hat{D} = D \cup \{\Delta\}$ be the one-point compactification of D. Thus Δ is identified with ∂D if D is bounded and with ∂D and the point at infinity if D is unbounded.

Define the explosion time $e : C([0,\infty); \hat{D}) \to [0,+\infty]$, $e(\omega) := \inf\{t \geqslant 0:\ \omega(t) = \Delta\}$ and introduce the space

$$\Omega_{\hat{D}} := \big\{\omega \in C\big([0,\infty); \hat{D}\big):\ \text{either } e(\omega) = \infty \text{ or} \\ \text{if } e(\omega) < \infty \text{ then } \omega\big(e(\omega) + t\big) = \Delta, \text{ for any } t \geqslant 0\big\}.$$

Following [20, p. 40], one constructs a natural stochastic basis $(\Omega_{\hat{D}}, \mathcal{F}, (\mathcal{F}_t))$. A *generalized martingale solution* for L on D (or an L-diffusion on D with absorbing boundary) is a family of probability measures $\mathbb{P}^x$ on $\Omega_{\hat{D}}$, $x \in \hat{D}$, such that $\mathbb{P}^x(\omega \in \Omega_{\hat{D}}:\ \omega(0) = x) = 1$ and, for each $f \in C_0^2(\mathring{D})$, one has that the process M_t^f, see (2.4), is an $\mathcal{F}_t$-martingale (here one uses the convention of setting $f(\Delta) = 0$, for any $f \in C_0^2(\mathring{D})$).

Similarly to (2.5) one defines couplings distributions on $\Omega_{\hat{D}} \times \Omega_{\hat{D}}$ for generalized martingale solutions on D.

3. Gradient estimates in $\mathbb{R}^d$

Let us write down our assumptions on L_0.

Hypothesis 3.1.

(i) The coefficients q_{ij} and b_i are continuous on $\mathbb{R}^d$; $V \geqslant 0$ is bounded and Borel on $\mathbb{R}^d$.
(ii) There exists $\lambda_0 > 0$ such that $\langle q(x)h, h\rangle \geqslant \lambda_0 |h|^2$, $x \in \mathbb{R}^d$, $h \in \mathbb{R}^d$.
(iii) The unique L-diffusion process does not explode, starting from any $x \in \mathbb{R}^d$.
(iv) There exists a nonnegative function $g \in C(0,+\infty)$ such that $\int_0^1 g(s)\,\mathrm{d}s < \infty$ and

$$\sup_{|x-y|=r} \frac{1}{r}\big\{\big\|\sigma(x) - \sigma(y)\big\|^2 + 2\big\langle b(x) - b(y), x - y\big\rangle\big\} \leqslant g(r), \tag{3.1}$$

for all $r > 0$, where $\sup \emptyset = 0$ by convention and $\sigma := \sqrt{q - \lambda_0 I}$ is the unique symmetric nonnegative definite matrix-valued function such that $\sigma^2 = q - \lambda_0 I$.

Remark 3.2. Let us comment on these hypotheses. Under (i), (ii) there exists a unique martingale solution for L, see [22, Section 10] or [20, Section 1.10]. Thus, the martingale problem for L on $\mathbb{R}^d$ is well posed and, in addition, the strong Markov property holds.

Condition (iv) generalizes substantially the standard condition that $g(r) := cr$, for some $c > 0$, implying the uniqueness and regularity of strong solutions to the associated SDEs. Note that under the assumption that $g(r) := cr$, for some $c > 0$, we can treat also potentials V which are globally Lipschitz, see Corollary 3.5.

If σ is constant, then (iv) holds if the drift term b is uniformly continuous on $\mathbb{R}^d$. Indeed, in this case, we can take as g the modulus of continuity of b, i.e., $g(r) = \omega_b(r) := \sup_{|x-y|\leqslant r} |b(x) - b(y)|$; note that g is continuous on $[0,\infty)$.

By Hypothesis 3.1 we can also deal with the nonlocal Lipschitz coefficients considered in [9].

It is remarkable that condition (iv) allows even to treat the case of unbounded *α-Hölder continuous coefficients* $q(x)$ and $b(x)$. This follows applying the next lemma and taking in (3.1)

$$g(s) := C\big(s^{2\alpha-1} + s^{\alpha}\big), \quad \alpha \in (0,1),$$

for a sufficiently large constant $C > 0$. Note that, possibly replacing λ_0 with a smaller positive constant, we can assume that $\sigma(x)$ is uniformly positive definite.

Lemma 3.3. *Let $\sigma(x) \in \mathcal{L}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$, be a symmetric matrix. Assume that, for some $\lambda > 0$, $\langle \sigma(x)h, h\rangle \geqslant \lambda |h|^2$, $x, h \in \mathbb{R}^d$. Let*

$$q(x) = \sigma(x)^2 + \lambda I, \quad x \in \mathbb{R}^d.$$

Then one has:

$$\big\|\sigma(x) - \sigma(y)\big\| \leqslant \frac{1}{2\sqrt{\lambda}} \big\|q(x) - q(y)\big\|, \quad x, y \in \mathbb{R}^d. \tag{3.2}$$

Proof. We will use the following elementary fact. Let A and B be two $d \times d$ symmetric positive matrices such that all eigenvalues of A and B are not less than $\lambda > 0$. One has:

$$\big\|\sqrt{A} - \sqrt{B}\,\big\| \leqslant \frac{1}{2\sqrt{\lambda}} \|A - B\|. \tag{3.3}$$

We provide a short proof for the sake of completeness. Note that $X := \sqrt{B} - \sqrt{A}$ solves the equation

$$X\big(-\sqrt{A}\big) + \big(-\sqrt{B}\big)X = -(B - A). \tag{3.4}$$

Since $-\sqrt{A}$ and $-\sqrt{B}$ are both stable matrices, adapting the classical method used to treat the Lyapunov equation $AX + XA^* = -I$, it is not difficult to solve (3.4). There exists a unique $d \times d$ matrix which solves (3.4) and this is given by

$$X = \int_0^\infty \mathrm{e}^{-t\sqrt{B}}(B - A)\mathrm{e}^{-t\sqrt{A}}\,\mathrm{d}t.$$

By this formula we get easily (3.3). The proof is complete. □

3.1. Construction of the coupling

Here, starting from L, we construct a family of coupling processes, with values in $\mathbb{R}^{2d}$, depending on $\varepsilon > 0$. This construction is not standard due to the fact that the coefficients q and b are only continuous. These coupling processes will allow us to obtain gradient estimates which will be independent of ε.

Let us fix $\varepsilon \in (0, 1)$ and define

$$\begin{aligned} &u(x,y) := (x-y)/|x-y|, \qquad u_\varepsilon(x,y) := u(x,y)\frac{k(|x-y|)}{k(|x-y|)+\varepsilon}, \\ &C_\varepsilon(x,y) := \lambda_0\big(I - 2u_\varepsilon(x,y)u_\varepsilon(x,y)^*\big) + \sigma(x)\sigma(y)^*, \quad x,y\in\mathbb{R}^d,\ x\neq y, \end{aligned} \tag{3.5}$$

with σ given in (3.1). Here $k \in C([0,\infty))$, such that $k(0)=0$, $k(s) \geqslant s^{1/4}$, $s \geqslant 0$, and

$$\int_0^1 \frac{g(s)}{k^2(s)}\,\mathrm{d}s < \infty, \tag{3.6}$$

where g is given in (3.1). Take for instance $k(s) = (\int_0^s g(r)\,\mathrm{d}r \vee s)^{1/4}$. Note that u_ε is continuous on $\mathbb{R}^{2d}$. The coupling processes will be generated by

$$\begin{aligned} L_\varepsilon(x,y) := \frac{1}{2}\sum_{i,j=1}^d \left\{ q_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + q_{ij}(y)\frac{\partial^2}{\partial y_i \partial y_j} + 2C_{ij}^\varepsilon(x,y)\frac{\partial^2}{\partial x_i \partial y_j} \right\} \\ + \sum_{i=1}^d \left\{ b_i(x)\frac{\partial}{\partial x_i} + b_i(y)\frac{\partial}{\partial y_i} \right\}, \quad x\neq y,\ \varepsilon>0, \end{aligned}$$

where $C_\varepsilon(x,y) = (C_{ij}^\varepsilon(x,y))$. Hence the $2d\times 2d$ diffusion matrix $Q_\varepsilon(x,y)$ of L_ε is given by

$$\begin{pmatrix} q(x) & \lambda_0(I-2u_\varepsilon(x,y)u_\varepsilon(x,y)^*) + \sigma(x)\sigma(y)^* \\ \lambda_0(I-2u_\varepsilon(x,y)u_\varepsilon(x,y)^*) + \sigma(y)\sigma(x)^* & q(y) \end{pmatrix}$$

and its drift term is $\tilde{b}(x,y) := (b(x), b(y))$ (remark that if $\sigma = 0$ and $\varepsilon = 0$, then L_ε is the generator of the coupling by reflection, while when $\lambda_0 = 0$ it reduces to the March coupling, see [5] for more details). To check that the matrix Q_ε is symmetric and nonnegative, one uses (ii) and that

$$\big|\big(I - 2u_\varepsilon(x,y)\,u_\varepsilon(x,y)^*\big)h\big| \leqslant |h|, \quad h\in\mathbb{R}^d,\ (x,y)\in\mathbb{R}^{2d},\ \varepsilon\in(0,1).$$

Thus, L_ε is a possibly degenerate elliptic second order operator on $\mathbb{R}^{2d}$ with continuous coefficients on $\mathbb{R}^{2d}$. Following [12, Theorems 2.2.IV and 2.3.IV] and using Euler approximations, one gets that there exists a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, such that for any $z = (x,y)$, there exists a process $Z_t^{z,\varepsilon}$ with values in $\hat{\mathbb{R}}^{2d}$ (the standard one point compactification of $\mathbb{R}^{2d}$) such that $Z_0^z = z$ a.s., and for each $F \in C_0^2(\mathbb{R}^{2d})$,

$$M_t^F := F\big(Z_t^{z,\varepsilon}\big) - \int_0^t L_\varepsilon F\big(Z_s^{z,\varepsilon}\big)\,\mathrm{d}s, \quad t\geqslant 0, \tag{3.7}$$

is an $\mathcal{F}_t$-martingale. Now, to simplify notation, we drop the dependence on z and ε for $Z_t^{z,\varepsilon}$, i.e., we set $Z_t = Z_t^{z,\varepsilon}$. Let $Z_t =: (X_t, Y_t)$. The explosion time of Z_t is

$$e = \lim_{n\uparrow\infty} \tau_n, \quad \tau_n := \inf\big\{t\geqslant 0\colon \big(|X_t| + |Y_t|\big) \geqslant n\big\}.$$

Since the marginal processes X_t and Y_t are both L-diffusion (starting from x and y, respectively), we have $e = \infty$ a.s., by (iii). Let us consider the coupling time T of Z_t, i.e.,

$$T := \inf\{t \geqslant 0:\ X_t = Y_t\},$$

where we set $\inf \emptyset = \infty$ by convention. This is an $\mathcal{F}_t$-stopping time; define a new process X'_t,

$$X'_t := \begin{cases} X_t, & t \leqslant T, \\ Y_t, & t > T. \end{cases} \tag{3.8}$$

We will show that

$$X_t \quad \text{and} \quad X'_t \quad \text{are equal in distribution.} \tag{3.9}$$

To this end, we first prove that X'_t is an L-diffusion. We write, for any $f \in C_0^2(\mathbb{R}^d)$,

$$\begin{aligned} f(X'_t) - \int_0^t Lf(X'_s)\,\mathrm{d}s &= \left(f(X'_{t\wedge T}) - \int_0^{t\wedge T} Lf(X'_s)\,\mathrm{d}s \right) \\ &\quad + \left(f(X'_t) - f(X'_{t\wedge T}) - \int_{t\wedge T}^t Lf(X'_s)\,\mathrm{d}s \right) \\ &= M_t^1 + M_t^2. \end{aligned} \tag{3.10}$$

We note that, for any $t \geqslant 0$, a.s.

$$M_t^2 = f(X'_t) - f(X'_{t\wedge T}) - \int_{t\wedge T}^t Lf(X'_s)\,\mathrm{d}s = f(Y_t) - f(Y_{t\wedge T}) - \int_{t\wedge T}^t Lf(Y_s)\,\mathrm{d}s$$

and $M_t^1 = f(X_{t\wedge T}) - \int_0^{t\wedge T} Lf(X_s)\,\mathrm{d}s$. It turns out that M_t^1 and M_t^2 are both martingales. Indeed, concerning M_t^1, letting $F(x, y) = f(x)$, $x, y \in \mathbb{R}^d$, we know that

$$f(X_t) - \int_0^t Lf(X_s)\,\mathrm{d}s = F(Z_t) - \int_0^t L_\varepsilon F(Z_s)\,\mathrm{d}s$$

is a local martingale; but in addition, since $f \in C_0^2(\mathbb{R}^d)$, $f(X_t) - \int_0^t Lf(X_s)\,\mathrm{d}s$ is also bounded. Hence $f(X_t) - \int_0^t Lf(X_s)\,\mathrm{d}s$ is an $\mathcal{F}_t$-martingale. By the Doob optional stopping theorem, we infer that M_t^1 is an $\mathcal{F}_t$-martingale as well. Similarly one proves that also M_t^2 is a martingale. Hence X'_t is an L-diffusion. Since weak uniqueness holds for L, [20, Chapter 1], we conclude that (3.9) holds. Define the coupling process $U_t = (X'_t, Y_t)$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and recall that $X'_t = Y_t$ a.s., for any $t \geqslant T$.

In this section the law of $U_t = U_t^{z,\varepsilon}$ on Ω^{2d} will be denoted by $\mathbb{P}_\varepsilon^{x,y}$, for any $z = (x, y) \in \mathbb{R}^{2d}$. $\mathbb{P}_\varepsilon^{x,y}$ will be our *coupling distribution*.

The marginal measures $\mathbb{P}^x$ and $\mathbb{P}^y$ of $\mathbb{P}_\varepsilon^{x,y}$ on Ω^d determine both L-diffusion processes, starting from x and y, respectively. The basic coupling property is

$$\begin{aligned}\mathbb{E}^x f(x_t) - \mathbb{E}^y f(x_t) &= \mathbb{E}_\varepsilon^{x,y}\big(f(x_t) - f(y_t)\big) \leqslant \mathbb{E}_\varepsilon^{x,y}\big|f(x_t) - f(y_t)\big| \\ &= \mathbb{E}_\varepsilon^{x,y}\big(\big|f(x_t) - f(y_t)\big|1_{\{T>t\}}\big) \leqslant \|f\|_\infty \mathbb{P}_\varepsilon^{x,y}(T > t), \qquad (3.11)\end{aligned}$$

$f \in \mathcal{B}_b^+(\mathbb{R}^d)$, where x_t and (x_t, y_t) denote the canonical processes (on Ω^d and Ω^{2d}, respectively) and T is the coupling time on Ω^{2d}.

3.2. Uniform gradient estimates

Here we establish gradient estimates when $D = \mathbb{R}^d$. We write P_t instead of $P_t^{\mathbb{R}^d}$.

Theorem 3.4. *Assume Hypothesis* 3.1. *Then, for any $t > 0$, $f \in \mathcal{B}_b(\mathbb{R}^d)$, the map $x \mapsto P_t f(x)$ is Lipschitz continuous on $\mathbb{R}^d$. Moreover, the following assertions hold.*

(a) *If $V = 0$, then setting*

$$c_t = \inf_{r>0}\left\{\frac{\int_0^r \exp[\frac{1}{4\lambda_0}\int_0^s g(u)\,\mathrm{d}u]\,\mathrm{d}s}{2\lambda_0 t} + \frac{1}{\int_0^r \exp\big[-\frac{1}{4\lambda_0}\int_0^s g(u)\,\mathrm{d}u\big]\,\mathrm{d}s}\right\}, \qquad (3.12)$$

one has:

$$\frac{\|\nabla P_t f\|_\infty}{\|f\|_\infty} \leqslant c_t, \quad t > 0,\ f \in \mathcal{B}_b^+\big(\mathbb{R}^d\big).$$

If, in particular, $C(\infty) := \int_0^\infty g(s)\,\mathrm{d}s < \infty$, then

$$\frac{\|\nabla P_t f\|_\infty}{\|f\|_\infty} \leqslant \frac{\exp[\frac{1}{4\lambda_0}C(\infty)](1 + 2\lambda_0)}{2\lambda_0\sqrt{t}}, \quad t > 0,\ f \in \mathcal{B}_b^+\big(\mathbb{R}^d\big). \qquad (3.13)$$

(b) *If $V \neq 0$, formula* (1.1) *holds with*

$$c := \big(1 + \|V\|_\infty\big)\frac{1 + 2\lambda_0}{2\lambda_0}\exp\left[\frac{1}{4\lambda_0}\int_0^1 g(s)\,\mathrm{d}s\right]. \qquad (3.14)$$

Proof. We will use the L_ε-diffusion previously constructed. Recall that $\rho(x, y) := |x - y|$ and, if no confusion may arise, we will write ρ instead of $|x - y|$. We need the following straightforward formulas (see [5, (2.8)])

$$L_\varepsilon(f \circ \rho) = \frac{1}{2}\bar{A}_\varepsilon f''(\rho) + \frac{f'(\rho)}{2\rho}(\operatorname{tr} A_\varepsilon - \bar{A}_\varepsilon + 2B), \quad f \in C^2(0, +\infty), \qquad (3.15)$$

where

$$A_\varepsilon := q(x) + q(y) - 2C_\varepsilon(x, y),$$
$$B := \langle b(x) - b(y), x - y \rangle,$$
$$\bar{A}_\varepsilon := \langle A_\varepsilon u(x, y), u(x, y) \rangle.$$

It is easy to check from (3.5) that

$$\bar{A}_\varepsilon \geqslant a_\varepsilon(\rho), \quad \text{where } a_\varepsilon(\rho) := 4\lambda_0 \frac{k^2(\rho)}{(k(\rho)+\varepsilon)^2}, \quad \text{and}$$
$$\operatorname{tr} A_\varepsilon = \|\sigma(x) - \sigma(y)\|^2 + a_\varepsilon(\rho), \quad \rho \geqslant 0.$$

It follows that

$$\operatorname{tr} A_\varepsilon - \bar{A}_\varepsilon + 2B \leqslant \|\sigma(x) - \sigma(y)\|^2 + 2\langle b(x) - b(y), x - y \rangle,$$

where $\|\sigma(x) - \sigma(y)\|^2 = \operatorname{tr}\{(\sigma(x) - \sigma(y))(\sigma(x) - \sigma(y))^*\}$. Thus, (3.1) and (3.15) imply that, for any $f \in C^2(0, \infty)$ with $f'' \leqslant 0$, $f' \geqslant 0$,

$$L_\varepsilon(f \circ \rho) \leqslant \frac{a_\varepsilon(\rho)}{2} f''(\rho) + \frac{1}{2} g(\rho) f'(\rho), \quad \rho > 0. \tag{3.16}$$

Now we split the proof into two parts.

Case I. $V = 0$. Let $a_0(r) = 4\lambda_0$, $r \geqslant 0$, and, for any $\varepsilon \geqslant 0$, define

$$\xi_\varepsilon(r) := \exp\left[\int_0^r \frac{g(s)}{a_\varepsilon(s)} \, \mathrm{d}s\right] = \exp\left[\frac{1}{4\lambda_0} \int_0^r \frac{g(s)(k(s)+\varepsilon)^2}{k^2(s)} \, \mathrm{d}s\right], \qquad h_\varepsilon(r) := \int_0^r \frac{\mathrm{d}s}{\xi_\varepsilon(s)},$$
$$r \geqslant 0. \tag{3.17}$$

Note that, by our hypothesis on the function k, ξ_ε is well defined. For any $\delta > 0$, set

$$F_{\delta,\varepsilon}(r) := \int_0^r \frac{1}{\xi_\varepsilon(s)} \left(\int_s^\delta \frac{\xi_\varepsilon(u)}{a_\varepsilon(u)} \, \mathrm{d}u \right) \mathrm{d}s, \quad r \in [0, \delta].$$

In this way one has

$$\frac{a_\varepsilon}{2} F''_{\delta,\varepsilon} + \frac{1}{2} g F'_{\delta,\varepsilon} = -\frac{1}{2},$$

$F'_{\delta,\varepsilon} \geqslant 0$ and $F''_{\delta,\varepsilon} \leqslant 0$ in $(0, \delta]$. By (3.16) we have:

$$L_\varepsilon(F_{\delta,\varepsilon} \circ \rho)(x, y) \leqslant -\frac{1}{2}, \quad \rho(x, y) \leqslant \delta, \ x \neq y.$$

Let $T_n := \inf\{t \geqslant 0: |x_t - y_t| < 1/n\}$ ($T_n \uparrow T$ as $n \to \infty$), and let

$$\begin{aligned} S_\delta &:= \inf\{t \geqslant 0\colon\ \rho(x_t, y_t) \geqslant \delta\}, \\ \sigma_N &:= \inf\{t \geqslant 0\colon\ |x_t| + |y_t| \geqslant N\}, \quad N \geqslant 1. \end{aligned} \tag{3.18}$$

Given $x \neq y$, choose $\delta > 0$ and $n_0 \geqslant 1$ such that $\delta > |x - y| > 1/n_0$. Then

$$0 \leqslant \mathbb{E}_\varepsilon^{x,y} F_{\delta,\varepsilon} \circ \rho(x_{T_n \wedge S_\delta \wedge \sigma_N}, y_{T_n \wedge S_\delta \wedge \sigma_N}) \leqslant F_{\delta,\varepsilon}\big(\rho(x, y)\big) - \frac{1}{2}\mathbb{E}_\varepsilon^{x,y} T_n \wedge S_\delta \wedge \sigma_N,$$

for any $n \geqslant n_0$ and $N > 1$. Since the L_ε-diffusion process is non-explosive, $\sigma_N \to \infty$ as $N \to \infty$, $\mathbb{P}_\varepsilon^{x,y}$ a.s. Therefore, letting $N \to \infty$, we get, by the above inequality,

$$\mathbb{E}_\varepsilon^{x,y} T_n \wedge S_\delta \leqslant 2F_{\delta,\varepsilon}\big(\rho(x, y)\big). \tag{3.19}$$

On the other hand, one has $L_\varepsilon(h_\varepsilon \circ \rho) \leqslant 0$, see (3.17), and so

$$h_\varepsilon\big(\rho(x, y)\big) \geqslant \mathbb{E}_\varepsilon^{x,y} h_\varepsilon \circ \rho(x_{T_n \wedge S_\delta \wedge \sigma_N}, y_{T_n \wedge S_\delta \wedge \sigma_N}) \geqslant h_\varepsilon(\delta)\mathbb{P}_\varepsilon^{x,y}(T_n \wedge \sigma_N > S_\delta).$$

Letting $N \to \infty$, we arrive at

$$\mathbb{P}_\varepsilon^{x,y}(T_n > S_\delta) \leqslant h_\varepsilon\big(\rho(x, y)\big)/h_\varepsilon(\delta).$$

Combining this with (3.19) and letting $n \to \infty$, we obtain

$$\begin{aligned} \mathbb{P}_\varepsilon^{x,y}(T > t) &= \mathbb{P}_\varepsilon^{x,y}(T > t, S_\delta > t) + \mathbb{P}_\varepsilon^{x,y}(T > t, S_\delta \leqslant t) \\ &\leqslant \mathbb{P}_\varepsilon^{x,y}(T \wedge S_\delta > t) + \mathbb{P}_\varepsilon^{x,y}(T > S_\delta) \\ &\leqslant \frac{2F_{\delta,\varepsilon}(\rho(x, y))}{t} + \frac{h_\varepsilon(\rho(x, y))}{h_\varepsilon(\delta)}, \quad t > 0,\ \rho(x, y) < \delta. \end{aligned} \tag{3.20}$$

Therefore, for any $f \in \mathcal{B}_b^+(\mathbb{R}^d)$, by the coupling property, see (3.11),

$$\begin{aligned} \frac{|P_t f(x) - P_t f(y)|}{|x - y|} &\leqslant \frac{\mathbb{E}_\varepsilon^{x,y}|f(x_t) - f(y_t)|}{|x - y|} \leqslant \frac{\mathbb{P}_\varepsilon^{x,y}(T > t)\|f\|_\infty}{|x - y|} \\ &\leqslant \left(\frac{2F_{\delta,\varepsilon}(|x - y|)}{t|x - y|} + \frac{h_\varepsilon(|x - y|)}{h_\varepsilon(\delta)|x - y|}\right)\|f\|_\infty \\ &\leqslant \left(\frac{2F'_{\delta,\varepsilon}(0)}{t} + \frac{h'_\varepsilon(0)}{h_\varepsilon(\delta)}\right)\|f\|_\infty, \quad \text{for } \rho(x, y) < \delta,\ t > 0. \end{aligned} \tag{3.21}$$

The last inequality follows since $F_{\delta,\varepsilon}$, $h_\varepsilon \in C^1([0, \delta])$, $\varepsilon \geqslant 0$, and $F'_{\delta,\varepsilon}$, h'_ε are both decreasing on $[0, \delta]$. Now, since $P_t f$ is bounded on $\mathbb{R}^d$, the previous formula shows that $P_t f$ is Lipschitz continuous on $\mathbb{R}^d$, for any $t > 0$. Indeed if $\rho(x, y) \geqslant \delta$ one has:

$$\big|P_t f(x) - P_t f(y)\big| \leqslant 2\|P_t f\|_\infty \leqslant \frac{2}{\delta}\|f\|_\infty \rho(x, y).$$

Now, letting $y \to x$ in (3.21), we arrive at

$$\limsup_{y \to x} \frac{|P_t f(x) - P_t f(y)|}{|x-y|} = \left|\nabla P_t f(x)\right| \leqslant \left(\frac{2F'_{\delta,\varepsilon}(0)}{t} + \frac{1}{h_\varepsilon(\delta)}\right)\|f\|_\infty,$$

for any $\varepsilon, \delta > 0$ (uniformly in $x \in \mathbb{R}^d$). Finally, letting $\varepsilon \to 0^+$, thanks to the properties of k, we get

$$\left|\nabla P_t f(x)\right| \leqslant \left(\frac{2F'_{\delta,0}(0)}{t} + \frac{1}{h_0(\delta)}\right)\|f\|_\infty, \quad \delta > 0.$$

By taking the infimum over all $\delta > 0$, we achieve (3.12). The desired c in (1.1) follows by taking $\delta = \sqrt{t \wedge 1}$. Finally, if $\int_0^\infty g(s)\,\mathrm{d}s < \infty$, then (3.13) follows by taking $\delta = \sqrt{t}$.

Case II. $V \neq 0$. We write

$$\begin{aligned}
\left|P_t f(x) - P_t f(y)\right| &= \left|\mathbb{E}_\varepsilon^{x,y}\left\{f(x_t)\mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s} - f(y_t)\mathrm{e}^{-\int_0^t V(y_s)\,\mathrm{d}s}\right\}\right| \\
&\leqslant \mathbb{E}_\varepsilon^{x,y}\left|\left(f(x_t) - f(y_t)\right)\mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s}\right| \\
&\quad + \|f\|_\infty \mathbb{E}_\varepsilon^{x,y}\left|\mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s} - \mathrm{e}^{-\int_0^t V(y_s)\,\mathrm{d}s}\right| \\
&\leqslant \|f\|_\infty\left(\mathbb{P}_\varepsilon^{x,y}(T > t) + \mathbb{E}_\varepsilon^{x,y}\left|\mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s} - \mathrm{e}^{-\int_0^t V(y_s)\,\mathrm{d}s}\right|\right).
\end{aligned} \tag{3.22}$$

Thus, to get the gradient estimate, we need to control the term

$$\begin{aligned}
\Lambda_{x,y}^\varepsilon &:= \mathbb{E}_\varepsilon^{x,y}\left|\mathrm{e}^{-\int_0^t V(x_s)\,\mathrm{d}s} - \mathrm{e}^{-\int_0^t V(y_s)\,\mathrm{d}s}\right| \leqslant \mathbb{E}_\varepsilon^{x,y}\left|\int_0^t V(x_s)\,\mathrm{d}s - \int_0^t V(y_s)\,\mathrm{d}s\right| \\
&\leqslant \int_0^t \mathbb{E}_\varepsilon^{x,y}\left|V(x_s) - V(y_s)\right|\mathrm{d}s \leqslant \|V\|_\infty \int_0^t \mathbb{P}_\varepsilon^{x,y}(T > s)\,\mathrm{d}s.
\end{aligned} \tag{3.23}$$

Now, using (3.20) one has, for $\rho(x,y) < \delta$,

$$\begin{aligned}
\int_0^t \mathbb{P}_\varepsilon^{x,y}(T > s)\,\mathrm{d}s &\leqslant \int_0^t \left(\mathbb{P}_\varepsilon^{x,y}(T \wedge S_\delta > s) + \mathbb{P}_\varepsilon^{x,y}(T > S_\delta)\right)\mathrm{d}s \\
&\leqslant \mathbb{E}_\varepsilon^{x,y}\int_0^t 1_{\{T \wedge S_\delta > s\}}\,\mathrm{d}s + t\frac{h_\varepsilon(\rho(x,y))}{h_\varepsilon(\delta)} \leqslant t\frac{h_\varepsilon(\rho(x,y))}{h_\varepsilon(\delta)} + \mathbb{E}_\varepsilon^{x,y}(T \wedge S_\delta \wedge t) \\
&\leqslant t\frac{h_\varepsilon(\rho(x,y))}{h_\varepsilon(\delta)} + 2F_{\delta,\varepsilon}\left(\rho(x,y)\right), \quad t > 0.
\end{aligned} \tag{3.24}$$

This implies, see (3.21),

$$\frac{\Lambda^{\varepsilon}_{x,y}}{|x-y|} \leqslant \|V\|_\infty \int_0^t \frac{\mathbb{P}^{x,y}_{\varepsilon}(T>s)}{|x-y|}\,\mathrm{d}s \leqslant \|V\|_\infty \left(\frac{t}{h_\varepsilon(\delta)} + 2F'_{\delta,\varepsilon}(0)\right), \tag{3.25}$$

for $\rho(x,y)<\delta$. This shows that $P_t f$ is Lipschitz continuous on $\mathbb{R}^d$, for any $t>0$.

Moreover, using (3.22) and letting $y\to x$, we get

$$\frac{|\nabla P_t f(x)|}{\|f\|_\infty} \leqslant \left(\frac{2F'_{\delta,\varepsilon}(0)}{t} + \frac{1}{h_\varepsilon(\delta)}\right) + \|V\|_\infty \left(\frac{t}{h_\varepsilon(\delta)} + 2F'_{\delta,\varepsilon}(0)\right).$$

Letting $\varepsilon\to 0^+$, we find, for any $t>0$,

$$\begin{aligned}\frac{|\nabla P_t f(x)|}{\|f\|_\infty} &\leqslant \inf_{\delta>0}\left(\frac{2F'_{\delta,0}(0)}{t} + \frac{1}{h_0(\delta)}\right)\big(1+t\|V\|_\infty\big)\\ &\leqslant \left(\frac{2F'_{\sqrt{t\wedge 1},0}(0)}{t} + \frac{1}{h_0(\sqrt{t\wedge 1})}\right)\big(1+t\|V\|_\infty\big).\end{aligned}$$

Now the estimate in (1.1), with c given in (3.14), follows using the semigroup property and the fact that $\|P_t f\|_\infty \leqslant \|f\|_\infty$, $f\in\mathcal{B}_b(\mathbb{R}^d)$, $t\geqslant 0$. □

Corollary 3.5. *If the function g in* (3.1) *verifies $g(r)\leqslant kr$, $r\geqslant 0$, for some $k>0$, then in Theorem* 3.4 *we can replace the assumption that V is bounded and Borel with the following one*:

V is Lipschitz continuous on $\mathbb{R}^d$.

Moreover, if C denotes the Lipschitz constant of V, we have (1.1) *with*

$$c := 1 + \frac{2C}{k}\big(\mathrm{e}^{k/2}-1\big) + \frac{1+2\lambda_0}{2\lambda_0}\mathrm{e}^{k/8\lambda_0}.$$

Proof. Note that the hypothesis on g guarantees that there is no explosion for the L-diffusion process. Since V is Lipschitz continuous, we have, see (3.23),

$$\Lambda^{\varepsilon}_{x,y} \leqslant \int_0^t \mathbb{E}^{x,y}_{\varepsilon}\big|V(x_s)-V(y_s)\big|1_{\{T>s\}}\,\mathrm{d}s \leqslant C\int_0^t \mathbb{E}^{x,y}_{\varepsilon}|x_s-y_s|1_{\{T>s\}}\,\mathrm{d}s.$$

If $g(r)\leqslant kr$, then (3.16) implies (with $f(s)=s$)

$$\mathbb{E}^{x,y}_{\varepsilon}|x_{s\wedge\sigma_N\wedge T_n} - y_{s\wedge\sigma_N\wedge T_n}| \leqslant |x-y| + \frac{k}{2}\int_0^s \mathbb{E}^{x,y}_{\varepsilon}|x_{r\wedge\sigma_N\wedge T_n} - y_{r\wedge\sigma_N\wedge T_n}|\,\mathrm{d}r.$$

Letting N and n tend to ∞ and using the Gronwall lemma, we infer

$$\mathbb{E}_\varepsilon^{x,y}|x_s - y_s|1_{\{T>s\}} \leqslant \mathbb{E}_\varepsilon^{x,y}|x_{s\wedge T} - y_{s\wedge T}| \leqslant |x-y|\mathrm{e}^{ks/2}, \quad s \geqslant 0,$$

for any $\varepsilon > 0$. Thus,

$$\limsup_{\varepsilon\to 0^+}\left(\limsup_{y\to x}\frac{\Lambda_{x,y}^\varepsilon}{|x-y|}\right) \leqslant C\int_0^t \mathrm{e}^{ks/2}\,\mathrm{d}s = \frac{2C}{k}\left(\mathrm{e}^{kt/2}-1\right).$$

Therefore, the desired result follows from the proof of Theorem 3.4. □

Theorem 3.4 leads to the following new Liouville theorem about space–time bounded harmonic functions. We refer to [1,7,21] and references therein for corresponding results concerning L-harmonic functions.

A bounded Borel function u on $[0,\infty)\times\mathbb{R}^d$ is called *space–time harmonic with respect to P_t* if

$$P_s u(t+s,\cdot)(x) = u(t,x), \quad s,t \geqslant 0,\ x\in\mathbb{R}^d.$$

It is trivial to see that any classical bounded solution to $\partial_t u + Lu = 0$ on $[0,\infty)\times\mathbb{R}^d$ is a space–time bounded harmonic function. Note that all bounded space–time harmonic functions of P_t are constant if and only if there exists a successful coupling for the L-diffusion process, see [7].

Theorem 3.6. *Assume that Hypothesis* 3.1 *holds with g satisfying*

$$\int_0^\infty \exp\left[-\frac{1}{4\lambda_0}\int_0^r g(s)\,\mathrm{d}s\right]\mathrm{d}r = \infty, \tag{3.26}$$

then any bounded space–time harmonic function u with respect to P_t is constant.

Proof. By (3.20), for $x,y\in\mathbb{R}^d$, $\delta > |x-y|$ and $t>0$, one has:

$$\begin{aligned}\left|u(t,x)-u(t,y)\right| &= \left|P_s u(t+s,x) - P_s u(t+s,y)\right| \leqslant \mathbb{E}_\varepsilon^{x,y}\left|u(t+s,x_s)-u(t+s,y_s)\right| \\ &\leqslant 2\|u\|_\infty \mathbb{P}_\varepsilon^{x,y}(T>s) \leqslant 2\left(\frac{2F_{\delta,\varepsilon}(|x-y|)}{s} + \frac{h_\varepsilon(|x-y|)}{h_\varepsilon(\delta)}\right)\|u\|_\infty.\end{aligned}$$

Letting first $\varepsilon\to 0^+$ and then $s\to\infty$ we get

$$\left|u(t,x)-u(t,y)\right| \leqslant 2\frac{h_0(|x-y|)}{h_0(\delta)}\|u\|_\infty.$$

Now if $\lim_{\delta\to\infty} h_0(\delta) = \infty$ (which is (3.26)), we get that $u(t,x) = u(t,y)$ and so $u(t,\cdot)$ is constant, denoted by u_t. Thus, since u is space–time harmonic, one has $u_0 = u(0,x) = P_t u(t,\cdot)(x) = u_t$, $t\geqslant 0$. Therefore, u is constant. □

Note that (3.26) holds in particular if $\int_0^\infty g(s)\,\mathrm{d}s < \infty$, compare with (3.13) in Theorem 3.4. As a consequence of Theorem 3.6, we now extend a result proved in [1].

Corollary 3.7. *Assume that conditions* (i)–(iii) *in Hypothesis* 3.1 *hold. Then any bounded space–time harmonic function with respect to P_t is constant if there exists $r > \frac{1}{4\lambda_0}$ such that*

$$r\|q(x) - q(y)\|^2 + 2\langle b(x) - b(y), x - y\rangle \leqslant 0, \quad x, y \in \mathbb{R}^d. \tag{3.27}$$

Proof. Since $r > \frac{1}{4\lambda_0}$, we have $\lambda_0 > \lambda' := \lambda_0 - \frac{1}{4r} > 0$. Then our assumptions hold for λ' in place of λ_0 with $\sigma := \sqrt{q - \lambda' I}$ bounded below by $\sqrt{\lambda_0 - \lambda'}\, I = \frac{1}{2\sqrt{r}} I$. By Lemma 3.3 and using (3.27) one obtains

$$\|\sigma(x) - \sigma(y)\|^2 + 2\langle b(x) - b(y), x - y\rangle \leqslant 0, \quad x, y \in \mathbb{R}^d. \tag{3.28}$$

Therefore, the desired assertion follows from Theorem 3.6 with $g := 0$. □

4. Gradient estimates in smooth domains

Here we assume that $D \neq \mathbb{R}^d$.

Hypothesis 4.1. We keep assumptions (i), (ii), (iv) of Hypothesis 3.1, replacing $\mathbb{R}^d$ with D. We add the following hypotheses:

(iii′) there exists a nonnegative function $\phi \in C^2(D)$ and a constant $\mu > 0$ such that

$$\lim_{r\to\infty} \inf_{x\in D,\,|x|\geqslant r} \phi(x) = \infty, \qquad L\phi(x) \leqslant \mu\phi(x), \quad x \in D. \tag{4.1}$$

(v) Let $\alpha > 0$ be such that $\rho_{\partial D}$ is smooth in D_α, see (2.2). There exist a positive constant β, a function $\gamma \in C^2(D; \mathbb{R}_+)$, strictly positive in $\mathring{D}$ and which vanishes on ∂D, and a nonnegative function $g_1 \in C(0, \alpha] \cap L^1(0, \alpha)$, such that if $0 < \rho_{\partial D}(x) < \alpha$ then

$$|\nabla\gamma(x)| \geqslant 1, \qquad \gamma(x) \leqslant \beta\rho_{\partial D}(x) \quad \text{and} \quad 2L\gamma(x) \leqslant g_1(\gamma(x)). \tag{4.2}$$

Remark 4.2. We note that (iii′) implies that the L-diffusion process is not explosive before hitting the boundary ∂D, see, e.g., [22, Section 10.2]. When D is bounded, (iii′) is trivial as $\inf\emptyset = \infty$ by convention.

Condition (v) generalizes the corresponding one in [10], namely, $L\rho_{\partial D} \leqslant M$ on $\{\rho_{\partial D} \leqslant \alpha\}$, for some constant $M > 0$. If this does not hold then gradient estimates can fail, see the counterexample in [25] and also [10]. To see that (v) strictly generalizes the condition used in [10], let us consider the next example.

Example 4.3. Consider $D := \{z = (x, y)\colon\ x \geqslant 0\} \subset \mathbb{R}^2$ and take the operator L such that $q(z) := I$, $z \in D$, and

$$b(x, y) := \big(y^+ \wedge (x^{-1/2}), 0\big), \quad b(0, y) := \big(y^+, 0\big), \quad x \geqslant 0,\ y \in \mathbb{R},$$

where $y^+ = (|y| + y)/2$. It is easy to verify that (i)–(iii) in Hypothesis 4.1 holds (note that b has at most linear growth and so (iii) is verified). However b is not even uniformly continuous on D. To check (iv), one considers

$$h(r) := \sup_{|z_1 - z_2| = r,\ z_1, z_2 \in D} \frac{\langle b(z_2) - b(z_1), z_2 - z_1 \rangle}{r}, \quad r > 0.$$

It is straightforward to get that $h(r) \leqslant r \vee r^{-1/2}$, $r > 0$; hence (iv) holds with $g(r) = r \vee r^{-1/2}$.

Concerning (v) one has: $\rho_{\partial D}(z) = x$ and $L\rho_{\partial D}(z) = y^+ \wedge x^{-1/2}$, $z = (x, y) \in D$, which is unbounded in any D_α, compare with [10]. However condition (v) holds with $\gamma := \rho_{\partial D}$ and $g_1(r) := r^{-1/2}$.

Now, similarly to Section 3.1, we construct a suitable coupling.

First, we extend q and b to the whole space $\mathbb{R}^d$ in the following way. With the help of the distance function ρ_∂, we extend these functions by reflection with respect to ∂D to $\tilde{D}_\alpha := \{x \notin D\colon \operatorname{dist}(x, D) < \alpha\}$. We still denote by q and b such extended coefficients. Then we take a continuous function $\psi : \mathbb{R}^d \to \mathbb{R}$, $0 \leqslant \psi \leqslant 1$, such that $\psi = 0$ on $\mathbb{R}^d \setminus (D \cup \tilde{D}_{\alpha/2})$ and $\psi = 1$ on D. Let us define:

$$\hat{q}(x) = \psi(x) q(x), \quad \hat{b}(x) = \psi(x) b(x), \quad x \in \mathbb{R}^d,$$

with the understanding that $q(x) = b(x) = 0$ if $x \notin D \cup \tilde{D}_\alpha$. Thus, in the sequel we assume that the operator L is defined on the whole space. Note that the extended L-diffusion process might be explosive. Next, we use the same notation of Section 3.1.

Let $z = (x, y) \in D \times D$. Set $\tilde{Z}_t = \tilde{Z}_t^{z,\varepsilon}$ be one L_ε-martingale solution with values in $\hat{\mathbb{R}}^{2d}$. Note that the $\tilde{e} > \tau_{D\times D}$, where $\tilde{e}$ is the explosion time of $\tilde{Z}_t =: (\tilde{X}_t, \tilde{Y}_t)$ and $\tau_{D\times D}$ is the hitting time to the boundary of $D \times D$. This fact follows by assumption (iii′) and by the marginality property. Remark that

$$\tau_{D\times D} = \tau_1 \wedge \tau_2, \quad \tau_1 = \inf\{t \geqslant 0\colon \tilde{X}_t \in \partial D\},\ \tau_2 = \inf\{t \geqslant 0\colon \tilde{Y}_t \in \partial D\}.$$

Let $\hat{D} = D \cup \{\Delta\}$ be the one-point compactification of D. Define a new process $Z_t = (X_t, Y_t)$, $t \geqslant 0$, as follows:

$$\begin{aligned} Z_t &:= \tilde{Z}_t, && \text{if } t < \tau_1 \wedge \tau_2,\\ Z_t &:= (\Delta, \tilde{Y}_t), && \text{if } \tau_1 \leqslant t < \tau_2,\\ Z_t &:= (\Delta, \Delta), && \text{if } t \geqslant \tau_2 \vee \tau_1,\\ Z_t &:= (\tilde{X}_t, \Delta), && \text{if } \tau_2 \leqslant t < \tau_1. \end{aligned}$$

It is clear that trajectories of Z_t are in $\Omega_{\hat{D}} \times \Omega_{\hat{D}}$. Now introduce a process X'_t as in (3.8). One shows that X_t and X'_t have the same distribution on $\Omega_{\hat{D}}$. To this end, one proceeds as in (3.10), replacing $f \in C_0^2(\mathbb{R}^d)$ with $f \in C_0^2(\mathring{D})$ and using Theorem 13.1 in [20, Chapter 1] about uniqueness of the generalized martingale solution for L on D.

Finally, we consider the coupling process (X'_t, Y_t), having law $\mathbb{P}_\varepsilon^{x,y}$ on $\Omega_{\hat{D}} \times \Omega_{\hat{D}}$. The measure $\mathbb{P}_\varepsilon^{x,y}$, $\varepsilon > 0$, $(x, y) \in D \times D$, will be our *coupling distribution*. The marginals of $\mathbb{P}_\varepsilon^{x,y}$ are $\mathbb{P}^x$ and $\mathbb{P}^y$, i.e., the generalized martingale solutions for L on D, starting from x and y, respectively.

Now, following the line of [25], we first estimate the hitting time to the boundary.

Lemma 4.4. *Under* (4.2) *there exists a constant* $c > 0$ *such that*

$$\mathbb{P}^x(\tau > t) \leqslant \frac{c\rho_{\partial D}(x)}{\sqrt{t \wedge 1}}, \quad x \in D,\ t > 0.$$

Proof. For any $f \in C^2((0,\alpha])$ we have

$$L(f \circ \gamma) = \frac{1}{2} f''(\gamma)\langle q\nabla\gamma, \nabla\gamma\rangle + f'(\gamma)L\gamma \quad \text{on } \{0 < \rho_{\partial D} < \alpha\}.$$

Similarly to the proof of Theorem 3.4, define

$$F_1(r) := \int_0^r \left(\mathrm{e}^{-\int_0^s g_1(u)/\lambda_0\,\mathrm{d}u} \int_s^\alpha \frac{\mathrm{e}^{\int_0^u g_1(u)/\lambda_0\,\mathrm{d}u}}{\lambda_0} \right) \mathrm{d}s, \qquad h_1(r) := \int_0^r \exp\left[-\int_0^s \frac{g_1(u)}{\lambda_0}\,\mathrm{d}u \right] \mathrm{d}s,$$

$r \in [0,\alpha]$, where g_1 is given in (4.2). Since $\langle q\nabla\gamma, \nabla\gamma\rangle \geqslant \lambda_0|\nabla\gamma|^2 \geqslant \lambda_0$ on $\{0 < \rho_{\partial D} < \alpha\}$, we have

$$L(F_1 \circ \gamma) \leqslant -\frac{1}{2} \quad \text{and} \quad L(h_1 \circ \gamma) \leqslant 0 \quad \text{on } \{0 < \rho_{\partial D} < \alpha\}.$$

Using that there is no explosion before hitting ∂D, the proof of Theorem 3.4 leading to (3.20), implies that, for $\rho_{\partial D}(x) < \alpha$,

$$\mathbb{P}^x(\tau > t) \leqslant \frac{2F_1(\gamma(x))}{t} + \frac{h_1(\gamma(x))}{h_1(\alpha)}, \quad t > 0.$$

Then

$$\mathbb{P}^x(\tau > t) \leqslant \frac{c\gamma(x)}{\sqrt{t \wedge 1}},$$

for some constant $c > 0$, and all $x \in D$ with $\rho_{\partial D}(x) < \alpha$. Thus, the desired assertion follows by noting that $\gamma(x) \leqslant \beta\rho_{\partial D}(x)$, for $\rho_{\partial D}(x) < \alpha$, and, for $\rho_{\partial D}(x) \geqslant \alpha$, one has

$$\mathbb{P}^x(\tau > t) \leqslant 1 \leqslant \frac{\rho_{\partial D}(x)}{\alpha\sqrt{t \wedge 1}}. \qquad \square$$

Combining the above lemma with the argument developed in Section 3, we are able to prove the following result.

Theorem 4.5. *Assume that Hypothesis* 4.1 *holds. Then, for any* $t > 0$, $f \in \mathcal{B}_b(\mathbb{R}^d)$, *the map* $x \mapsto P_t f(x)$ *is* (*globally*) *Lipschitz continuous on* D. *Moreover the uniform estimate* (1.1) *holds for some* $c > 0$.

Proof. Let $\mathbb{P}_\varepsilon^{x,y}$, $x, y \in D$, be the coupling distribution on $\Omega_{\hat{D}} \times \Omega_{\hat{D}}$ previously constructed. Let

$$\tau_1 := \inf\{t \geqslant 0:\ x_t \in \partial D\}, \qquad \tau_2 := \inf\{t \geqslant 0:\ y_t \in \partial D\}$$

and $T := \inf\{t \geqslant 0:\ x_t = y_t\}$ be the coupling time.

Case I. $V=0$. Now we split the proof into three parts.

(a) Let $z_t=(x_t,y_t)$ be the canonical coupling process. For any $t>0$, $x,\ y\in D,\ f\in\mathscr{B}_b^+(D)$, we have, see also [25],

$$\begin{aligned}\big|P_t^D f(x)-P_t^D f(y)\big| &= \big|\mathbb{E}^x f(x_t)1_{\{t<\tau\}}-\mathbb{E}^y f(x_t)1_{\{t<\tau\}}\big|\\ &\leqslant \mathbb{E}_\varepsilon^{x,y}\big|f(x_t)1_{\{t<\tau_1\}}-f(y_t)1_{\{t<\tau_2\}}\big|\\ &\leqslant \|f\|_\infty \mathbb{P}_\varepsilon^{x,y}(t<\tau_1\vee\tau_2, T>t\wedge\tau_1\wedge\tau_2)\\ &\leqslant \|f\|_\infty\big(\mathbb{P}_\varepsilon^{x,y}(t<\tau_1\vee\tau_2, T>t\wedge\tau_1\wedge\tau_2, T\wedge\tau_1\wedge\tau_2>t/2)\\ &\quad+\mathbb{P}_\varepsilon^{x,y}(t<\tau_1\vee\tau_2, T>t\wedge\tau_1\wedge\tau_2, T\wedge\tau_1\wedge\tau_2\leqslant t/2)\big).\end{aligned}$$

Now let $\rho(x,y)<\delta$ and S_δ be defined as in (3.18); we find

$$\begin{aligned}&\{t<\tau_1\vee\tau_2, T>t\wedge\tau_1\wedge\tau_2, T\wedge\tau_1\wedge\tau_2\leqslant t/2\}\\ &\quad\subset\{T\wedge\tau_1\wedge\tau_2>S_\delta\}\cup\left\{\tau_1\leqslant\frac{t}{2}\wedge S_\delta, t<\tau_2\right\}\cup\left\{\tau_2\leqslant\frac{t}{2}\wedge S_\delta, t<\tau_1\right\}.\end{aligned}$$

Hence

$$\begin{aligned}&\big|P_t^D f(x)-P_t^D f(y)\big|\\ &\quad\leqslant\|f\|_\infty\big\{\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2>t/2)+\mathbb{P}_\varepsilon^{x,y}\big(\tau_1\leqslant(t/2)\wedge S_\delta, \tau_2>t\big)\\ &\qquad+\mathbb{P}_\varepsilon^{x,y}\big(\tau_2\leqslant(t/2)\wedge S_\delta, \tau_1>t\big)+\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2>S_\delta)\big\},\quad |x-y|<\delta.\end{aligned}\tag{4.3}$$

Repeating the proof of Theorem 3.4, with T replaced by $T\wedge\tau_1\wedge\tau_2$, we obtain

$$\begin{aligned}&\big\{\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2>t/2)+\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2\geqslant S_\delta)\big\}\\ &\quad\leqslant\left(\frac{2F_{\delta,\varepsilon}(\rho(x,y))}{t}+\frac{h_\varepsilon(\rho(x,y))}{h_\varepsilon(\delta)}\right).\end{aligned}\tag{4.4}$$

Now we claim that

$$\mathbb{P}_\varepsilon^{x,y}\big(\tau_i\leqslant S_\delta\wedge(t/2), \tau_j>t\big)\leqslant\frac{c}{\sqrt{t\wedge 1}}\rho(x,y),\quad 1\leqslant i\neq j\leqslant 2,\tag{4.5}$$

for some constant $c>0$, and all $t>0$. Once (4.5) is proved, arguing as in (3.21), we find that (4.3)–(4.5) imply that $x\mapsto P_t f(x)$ is Lipschitz continuous on D. Moreover, we get:

$$\begin{aligned}\frac{|\nabla P_t f(x)|}{\|f\|_\infty}&\leqslant\limsup_{y\to x}\frac{1}{|x-y|}\big\{\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2>t/2)\\ &\quad+\mathbb{P}_\varepsilon^{x,y}\big(\tau_1\leqslant S_\delta\wedge(t/2), \tau_2>t\big)+\mathbb{P}_\varepsilon^{x,y}\big(\tau_2\leqslant S_\delta\wedge(t/2), \tau_1>t\big)\\ &\quad+\mathbb{P}_\varepsilon^{x,y}(T\wedge\tau_1\wedge\tau_2\geqslant S_\delta)\big\}\leqslant\frac{c}{\sqrt{t\wedge 1}},\quad t>0.\end{aligned}\tag{4.6}$$

Thus it remains to prove (4.5).

(b) We claim that, for any $\varepsilon > 0$, $x, y \in D$,

$$1_{\{\tau_1 \leqslant t\wedge\tau_2\}}\mathbb{E}_\varepsilon^{x,y}\big(1_{\{\tau_2-(\tau_1\wedge t)>t/2\}}|\mathcal{F}_{\tau_1\wedge t}\big) = 1_{\{\tau_1 \leqslant t\wedge\tau_2\}}\mathbb{P}^{y_{\tau_1\wedge t}}(\tau_2 > t/2), \quad \mathbb{P}_\varepsilon^{x,y} \text{ a.s.} \tag{4.7}$$

Note that to prove (4.7), we cannot use the strong Markov property for the coupling process, since in general this does not hold.

To simplify notation, set $\hat\tau_1 = \tau_1 \wedge t$. Introduce $\theta_{\hat\tau_1}(\omega)(t) = \omega(t + \hat\tau_1(\omega))$, $t \geqslant 0$, $\omega \in \Omega_{\hat D} \times \Omega_{\hat D}$. We have $\tau_2 = \tau_2 \circ \theta_{\hat\tau_1} + \hat\tau_1$ provided $\hat\tau_1 \leqslant \tau_2$. Thus, since $\{\hat\tau_1 \leqslant \tau_2\} \in \mathcal{F}_{\hat\tau_1}$, there holds

$$\begin{aligned}1_{\{\tau_1 \leqslant t\wedge\tau_2\}}\mathbb{E}_\varepsilon^{x,y}\big(1_{\{\tau_2-\hat\tau_1>t/2\}}|\mathcal{F}_{\hat\tau_1}\big) &= 1_{\{\tau_1 \leqslant t\wedge\tau_2\}}\mathbb{P}_\varepsilon^{x,y}\big(\tau_2\circ\theta_{\hat\tau_1} > t/2|\mathcal{F}_{\hat\tau_1}\big)\\ &= 1_{\{\tau_1 \leqslant t\wedge\tau_2\}}\mathbb{P}_\varepsilon^{x,y}\big(\theta_{\hat\tau_1}^{-1}\{\tau_2 > t/2\}|\mathcal{F}_{\hat\tau_1}\big).\end{aligned} \tag{4.8}$$

Since $\hat\tau_1 = \tau_1 \wedge t$ is a bounded stopping time and since the coefficients of L_ε are locally bounded, by [13, Lemma 5.4.19, p. 321] we have

$$\mathbb{P}_\varepsilon^{x,y}\big(\theta_{\tau_1\wedge t}^{-1}\{\tau_2 > t/2\}|\mathcal{F}_{\tau_1\wedge t}\big)(\omega) = \tilde{\mathbb{P}}_\varepsilon^{\omega}(\tau_2 > t/2), \tag{4.9}$$

where, for each ω, $\tilde{\mathbb{P}}_\varepsilon^{\omega}$ is one (not necessarily unique) solution to the L_ε-martingale problem, starting from $(x_{\tau_1\wedge t}(\omega), y_{\tau_1\wedge t}(\omega))$ (thus we could write $\tilde{\mathbb{P}}_\varepsilon^{\omega} = \tilde{\mathbb{P}}_\varepsilon^{x_{\tau_1}(\omega), y_{\tau_1}(\omega)}$ if $\tau_1 \leqslant t$). Because $\{\tau_2 > t/2\}$ depends only on the second component of ω, by the marginality (cf. (2.5)) one has:

$$\tilde{\mathbb{P}}_\varepsilon^{\omega}(\tau_2 > t/2) = \tilde{\mathbb{P}}_\varepsilon^{\omega}\circ\pi_2^{-1}(\tau_2 > t/2) = \mathbb{P}^{y_{\tau_1(\omega)\wedge t}}(\tau_2 > t/2), \quad \mathbb{P}_\varepsilon^{x,y} \text{ a.s.}$$

Combining this with (4.8) and (4.9) we prove (4.7).

(c) Here we show (4.5). By (4.7) and applying Lemma 4.4, we obtain

$$\begin{aligned}\mathbb{P}_\varepsilon^{x,y}\big(\tau_1 \leqslant S_\delta \wedge (t/2), \tau_2 > t\big) &= \mathbb{E}_\varepsilon^{x,y}1_{\{\tau_1 \leqslant (t/2)\wedge\tau_2\wedge S_\delta\}}1_{\{\tau_2>t\}}\\ &\leqslant \mathbb{E}_\varepsilon^{x,y}\big[\mathbb{E}_\varepsilon^{x,y}\big(1_{\{\tau_1 \leqslant (t/2)\wedge\tau_2\wedge S_\delta\}}1_{\{\tau_2-\tau_1\wedge t>t/2\}}|\mathcal{F}_{\tau_1\wedge t}\big)\big]\\ &= \mathbb{E}_\varepsilon^{x,y}1_{\{\tau_1 \leqslant (t/2)\wedge\tau_2\wedge S_\delta\}}\big[\mathbb{E}_\varepsilon^{x,y}\big(1_{\{\tau_2-\tau_1\wedge t>t/2\}}|\mathcal{F}_{\tau_1\wedge t}\big)\big]\\ &= \mathbb{E}_\varepsilon^{x,y}1_{\{\tau_1 \leqslant (t/2)\wedge\tau_2\wedge S_\delta\}}\mathbb{P}^{y_{\tau_1}}(\tau_2 > t/2)\\ &\leqslant \frac{c}{\sqrt{t\wedge 1}}\mathbb{E}_\varepsilon^{x,y}\rho_{\partial D}(y_{\tau_1})1_{\{\tau_1 \leqslant (t/2)\wedge\tau_2\wedge S_\delta\}}\\ &\leqslant \frac{c}{\sqrt{t\wedge 1}}\mathbb{E}_\varepsilon^{x,y}\rho\big(x_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}, y_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}\big),\end{aligned} \tag{4.10}$$

for some constant $c > 0$. Since $L_\varepsilon(h_\varepsilon \circ \rho) \leqslant 0$ on $D \times D$ and $c_1\rho \leqslant h_\varepsilon(\rho) \leqslant c_2\rho$ for some constants $c_1, c_2 > 0$, independent of ε, and all $\rho \leqslant \delta$, we have

$$\begin{aligned}\mathbb{E}_\varepsilon^{x,y}\rho\big(x_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}, y_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}\big) &\leqslant \frac{1}{c_1}\mathbb{E}_\varepsilon^{x,y}h_\varepsilon\circ\rho\big(x_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}, y_{\tau_1\wedge\tau_2\wedge S_\delta\wedge(t/2)}\big)\\ &\leqslant \frac{h_\varepsilon\circ\rho(x,y)}{c_1} \leqslant \frac{c_2}{c_1}\rho(x,y).\end{aligned}$$

Substituting this into (4.10) we obtain

$$\mathbb{P}_\varepsilon^{x,y}\big(\tau_1 \leqslant S_\delta \wedge (t/2), \tau_2 > t\big) \leqslant \frac{c}{\sqrt{t \wedge 1}}\rho(x, y),$$

for some constant $c > 0$. Similarly, the same holds by exchanging τ_1 and τ_2. Therefore, (4.5) holds.

Case II. $V \neq 0$. Using the gradient estimates already known for $V = 0$, the proof is similar to the one of Theorem 3.4 in the case of $V \neq 0$. We only remark that to treat the term

$$\int_0^t \mathbb{E}_\varepsilon^{x,y}\big|V(x_s)1_{\{s<\tau_1\}} - V(y_s)1_{\{s<\tau_2\}}\big|\,\mathrm{d}s,$$

one first uses estimate (4.3) (thanks to the boundedness of V). Then one proceeds as in (3.24) and (3.25), with T replaced by $T \wedge \tau_1 \wedge \tau_2$. In addition, one uses (4.5). The proof is complete. □

In particular, Theorem 4.5 implies the following classical gradient estimates for elliptic diffusion semigroups on compact regular domains.

Corollary 4.6. *If D is bounded, L is uniformly elliptic on D, $V = 0$ and q, b are Hölder continuous, then, for some $c, \lambda > 0$,*

$$\big\|\nabla P_t^D f\big\|_\infty \leqslant \frac{c\,\mathrm{e}^{-\lambda t}}{\sqrt{t}}\|f\|_\infty, \quad f \in \mathcal{B}_b^+(D),\ t > 0.$$

Proof. Let u be the positive first Dirichlet eigenfunction of Δ on D, one has $u|_{\partial D} = 0, u > 0$ in $\mathring{D}$ and $|\nabla u| > 0$ in a neighborhood of ∂D. Since D is compact and q is uniformly positive definite on D, Hypothesis 4.1 holds. In particular, (4.2) holds for some $\alpha > 0$, with $\gamma := Ru$ and $R > 0$ such that $|\nabla\gamma| \geqslant 1$ on $\{\rho_{\partial D} < \alpha\}$. Then Theorem 4.5 implies

$$\big\|\nabla P_t^D f\big\|_\infty \leqslant \frac{c}{\sqrt{t \wedge 1}}\big\|P_{t/2}^D f\big\|_\infty, \quad t > 0,\ f \in \mathcal{B}_b^+(D).$$

Since D is compact, one has:

$$\big\|P_{t/2}^D f\big\|_\infty \leqslant \|f\|_\infty \sup_{x\in D}\mathbb{P}^x(\tau > t/2) \leqslant c\mathrm{e}^{-\lambda t}\|f\|_\infty,$$

for some constants $c, \lambda > 0$. The proof is complete. □

Acknowledgments

The authors thank the Mathematics Departments of the Universities of Bielefeld, Torino and Poitiers, where parts of this paper were prepared, for good working conditions.

References

[1] M. Bertoldi, S. Fornaro, Gradient estimates in parabolic problems with unbounded coefficients, Studia Math. 165 (2004) 221–254.

[2] M. Bertoldi, L. Lorenzi, Estimates for the derivatives for parabolic operators with unbounded coefficients, Trans. Amer. Math. Soc. 357 (2005) 2627–2664.

[3] M. Bertoldi, S. Fornaro, L. Lorenzi, Gradient estimates for parabolic problems with unbounded coefficients in nonconvex domains, preprint, Tübinger Berichte zur Funktionalanalysis, 2004.

[4] S. Cerrai, Second Order PDE's in Finite and Infinite Dimension. A Probabilistic Approach, Lecture Notes in Math., vol. 1762, Springer, Berlin, 2001.

[5] M.-F. Chen, S.-F. Li, Coupling methods for multi-dimensional diffusion process, Ann. Probab. 17 (1989) 151–177.

[6] M. Cranston, A probabilistic approach to gradient estimates, Canad. Math. Bull. 35 (1992) 46–55.

[7] M. Cranston, A. Greven, Coupling and harmonic functions in the case of continuous time Markov processes, Stochastic Process. Appl. 60 (1995) 261–286.

[8] G. Da Prato, A. Lunardi, On the Ornstein–Uhlenbeck operator in spaces of continuous functions, J. Funct. Anal. 131 (1995) 94–114.

[9] S. Fang, T. Zhang, A class of stochastic differential equations with non-Lipschitzian coefficients: pathwise uniqueness and no explosion, C. R. Math. Acad. Sci. Paris 337 (2003) 737–740.

[10] S. Fornaro, G. Metafune, E. Priola, Gradient estimates for Dirichlet parabolic problems in unbounded domains, J. Differential Equations 205 (2004) 329–353.

[11] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, second ed., Springer, Berlin, 1983.

[12] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, first ed., North-Holland, Amsterdam, 1981.

[13] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, second ed., Springer, Berlin, 1998.

[14] M. Ledoux, A simple proof of an inequality by P. Buser, Proc. Amer. Math. Soc. 121 (1994) 951–958.

[15] T. Lindvall, L.C.G. Rogers, Coupling of multidimensional diffusions by reflection, Ann. Probab. 14 (1986) 460–872.

[16] T. Lindvall, Lectures on the Coupling Method, Corrected reprint of the 1992 original, Dover, Mineola, NY, 2002.

[17] A. Lunardi, Analytic Semigroups and Optimal Regularity in Parabolic Problems, Birkhäuser, Basel, 1995.

[18] A. Lunardi, Schauder theorems for linear elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^N$, Studia Math. 128 (1998) 171–198.

[19] G. Metafune, D. Pallara, M. Wacker, Feller semigroups on $\mathbb{R}^n$, Semigroup Forum 65 (2002) 159–205.

[20] R.G. Pinsky, Positive Harmonic Functions and Diffusion, Cambridge Stud. Adv. Math., vol. 45, Cambridge Univ. Press, Cambridge, 1995.

[21] E. Priola, J. Zabczyk, Liouville theorems for non-local operators, J. Funct. Anal. 216 (2004) 455–490.

[22] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, Springer, Berlin, 1979.

[23] F.-Y. Wang, Gradient estimates on $\mathbb{R}^d$, Canad. Math. Bull. 37 (1994) 560–570.

[24] F.-Y. Wang, Application of coupling method to the Neumann eigenvalue problem, Probab. Theory Related Fields 98 (1994) 299–306.

[25] F.-Y. Wang, Gradient estimates of Dirichlet heat semigroups and application to isoperimetric inequalities, Ann. Probab. 32 (2004) 424–440.